

Psychological Monographs

General and Applied

No. 353
1953

M. J. Freeman

The Development of a Test for the
Measurement of Anxiety: A Study
of Its Reliability and Validity

By

M. J. Freeman

North Hollywood, California

Vol. 67
No. 3

Price, \$1.00



Edited by Herbert S. Conrad
Published by The American Psychological Association, Inc.

Psychological Monographs: General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology*
with the *Psychological Monographs*

Editor

HERBERT S. CONRAD

Federal Security Agency
Office of Education
Washington 25, D.C.

Managing Editor

LORRAINE BOUTHILET

Consulting Editors

DONALD E. BAIER

FRANK A. BEACH

ROBERT G. BERNREUTER

WILLIAM A. BROWNELL

HAROLD E. BURTT

JERRY W. CARTER, JR.

CLYDE H. COOMBS

JOHN G. DARLEY

JOHN F. DASHIELL

EUGENIA HANFMANN

EDNA HEIDBREDER

HAROLD E. JONES

DONALD W. MACKINNON

LORRIN A. RIGGS

CARL R. ROGERS

SAUL ROSENZWEIG

ROSS STAGNER

PERCIVAL M. SYMONDS

JOSEPH TIFFIN

LEDYARD R. TUCKER

JOSEPH ZUBIN

MANUSCRIPTS should be sent to the Editor.

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded* or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

The first page of the manuscript should contain the title of the paper, the author's name, and his institutional connection (or his city of residence). Acknowledgments should be kept brief, and appear as a *footnote* on the first page. No table of contents need be included. For other directions or suggestions on the preparation of manuscripts, see: CONRAD, H. S. Preparation of manuscripts for publication as monographs. *J. Psychol.*, 1948, 26, 447-459.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as author's fees, subscriptions and sales, change of address, etc.) should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C.

COPYRIGHT, 1953, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Psychological Monographs: General and Applied

The Development of a Test for the Measurement of Anxiety: A Study of Its Reliability and Validity^{1, 2}

M. J. Freeman
North Hollywood, California

A review of the literature on anxiety reveals agreement on the fundamental assumption that fear, frustration, conflict, threat, and apprehension are basic facets of anxiety. Although in many instances a divergence of opinion exists as to the relative importance and primacy of these emotional forces, there is no dispute as to the centrality of anxiety in relation to the development of neurosis.

A. DEFINITION OF MANIFEST ANXIETY ("MA")

In this study the term *manifest anxiety* embodies the behavioral responses of

fear and apprehension which psychiatrists have observed to be manifest to an abnormal degree in persons suffering from neurosis.³ It was assumed that these observable manifestations are in effect cues to a nuclear syndrome of anxiety.

This study assumes that anxiety as a pathological manifestation is to be differentiated from normal anxiety which occurs in response to an actual external threat. The following psychiatric requirements for the diagnosis of manifest anxiety were laid down in the present study:

- a. Minimum observation period of 30 days by attending physicians, most of whom were psychiatrists.
- b. In each instance, medical findings which reported no organic pathology to account for the symptoms.
- c. In each instance, concurrence by the head hospital psychiatrist in the diagnosis of manifest anxiety.

B. SPECIFIC PURPOSES OF THE PRESENT INVESTIGATION

If the assumptions in the above sec-

¹ This test is now combined with the psychosomatic test which was published separately (3). Since the latter instrument was standardized upon the same subjects as in the present investigation, the final instrument provides a suitable measure of both anxiety neurosis and psychosomatic syndromes.

² In view of the fact that this final measuring instrument is based upon an original doctoral research performed at the Claremont Graduate School, as well as upon earlier graduate research at the University of Southern California, the author desires to express grateful acknowledgment to Dr. J. P. Guilford of the University of Southern California and to Dr. Florence Mateer, Dr. F. Theodore Perkins, and Dr. J. E. Caster of Claremont Graduate School, for their respective guidance in the various phases of this research. Thanks are also extended to Dr. Eugene Ziskind, Dr. Harry Roback, and Dr. Karl O. Von Hagen for their excellent cooperation in providing medical facilities for validation of the criterion cases.

³ The term "manifest" has been selected in order to identify that form of anxiety which is observable, and therefore capable of being identified, by the attending psychiatrist. It is to be distinguished from "latent" anxiety, which would provide a difficult basis upon which to make an objective evaluation. The term "manifest anxiety" therefore involves an overt expression of the inner (nuclear) feelings of abnormal fears and apprehension.

tion are correct, it follows that manifest anxiety, as defined, will not appear in a normal population, and that manifest anxiety will be present in the various forms of anxiety neuroses.

The problem here is twofold: (a) to develop a valid measuring instrument of manifest anxiety ("MA"), and (b) to differentiate "MA" cases from the normal population.

Specific questions for investigation were as follows:

1. Does the factor of hospitalization have an effect upon the development of manifest anxiety ("MA")?
2. Does sex have any relationship to "MA"?
3. What type of item would prove most successful in differentiating the "MA" criterion cases from the normals? What implications may be drawn from the content of the valid items?
4. Is the psychiatric classification of manifest anxiety a clinical entity capable of objective measurement?
5. Is manifest anxiety nuclear to all forms of psychoneurosis?

C. DEVELOPMENT OF TEST MATERIALS

The following factors have been given careful consideration in order to overcome the common weaknesses of a paper-and-pencil personality test:

- a. The elimination of questions or declarative items the answers to which an uncooperative subject could readily falsify; e.g., "Are you nervous?" and "I daydream very little."
- b. The construction of items with meaning-tones sufficiently disguised that the subject will be unaware of the diagnostic intent of the items. (This purpose, it is realized, cannot be fully accomplished in every instance.)
- c. The formulation of items of a projective character. The test is to bring

about projections through the respondent's use of the mechanism of identification. Examples:

Usually when there is pain in one part of the body, other parts of the body also become affected with pain.

() Yes () No

The person who presents the more undesirable behavior is

- (1) the one who criticizes people
- (2) the person who lies occasionally

() 1 () 2

A person who prefers to be alone is usually

- (1) an unfriendly person
- (2) one who minds his own business
- (3) a home lover

() 1 () 2 () 3

In the following type of item, the respondent answers in accordance with the following instructions:

Below is a list of paired characteristics. You are to place an "X" before the item of a pair which comes closest to fitting the behavior of the average person. Be sure that you mark one of each set of statements.

- () fears another war taking place
- () fears doing the wrong thing

Examination of the items above shows that the subject, in giving answers, would presumably need to respond through a projective mechanism of one sort or another.

D. TITLE AND GENERAL DESCRIPTION

The title of the complete measuring instrument is the *Freeman "MA" and "PS" Test*. "MA," of course, stands for Manifest Anxiety; "PS" for Psychosomatic. The "MA" test will be found in the appendix.

It will be observed that the instructions on the face sheet give the subject the impression that the purpose of the "MA" test is to determine his ability to judge the behavior of other people. The purpose of this disguise is to remove all possible tension and defensiveness in the subject in taking the test.

The "MA" part of the test contains five subsections listed as A, B, C, D, and E. There is a total of 141 items in these combined sections.

E. TECHNIQUE OF ADMINISTRATION

In an effort to standardize the "MA" test in realistic situations, the test was administered only under conditions which led the subject to regard the taking of the test as a serious performance.

Hospital and nonhospital subjects were given the test singly or in small groups. In some cases, the tests were given by the head psychiatrist or his assistant, while in other cases they were administered by the investigator, professor, or by the personnel manager. However, great care was taken in each instance to provide uniform instructions to the subjects.

In the hospital situations, the patients were told that the test was being given because the doctors needed the information to assist them in making a further diagnosis. In industrial situations, the test was administered to actual job applicants. In the college classrooms the professors of the respective classes gave the tests to their students with the instructions that the test results would be used by the professor for his own evaluation of the student.

Quite often after starting the test, the subjects would approach the examiner and ask the question: "How should I take the test, the way I think others feel or the way I feel?" In each instance, the answer was given: "It does not matter, either way you please." As indicated previously, it is assumed that the subject has no basis for response other than his own experiences.

Where subjects sought to obtain clarification of a particular question in order to be assisted in giving their best answer, they were simply again referred to the preceding instructions. No other information was provided.

It was required that the signature of each subject appear on the test.

F. SUBJECTS

In the development and validation of the test, 921 subjects were tested. It has taken a period of three years to develop the final measuring instrument.

The investigation was concerned with two principal groups: (a) Control ("Normals"), and (b) Criterion ("Manifest

Anxiety Cases").⁴ The two groups will hereafter be designated as the "N" and "MA" groups.

1. *The Criterion ("MA") Group*

Considerable attention has been devoted to the matter of obtaining a valid criterion. In keeping with the objective of avoiding any possible misconception regarding the character of manifest anxiety itself, every psychiatrist and resident physician was provided with the definition of manifest anxiety. In no instance did this investigator encounter any dispute regarding the existence and nature of this psychopathological trait.

Furthermore, the physicians were advised that only those cases in which manifest anxiety was severe were to be selected, i.e., termed as "Manifest Anxiety" cases. The classification of personality disorder was not of any importance for this study. Consideration was given to one factor only: the patients selected were to be those who, on the basis of medical observation, displayed signs of manifest anxiety in accordance with the definition of that term.

To further insure the validity of the criterion, sampling was performed not in one particular hospital as is so often the case, but in several different institutions. It is believed that this procedure eliminates the weakness of one-sided diagnosis, and furthermore it would seem to effect a more representative sampling of the criterion cases.

The sources of cases in the hospitals were not limited to the neuropsychiatric wards. A large percentage of the "N" cases was selected from gastrointestinal wards. Here the procedure for selection was as follows: the chief psychiatrist and the investigator visited the general wards

⁴The subgroups will be discussed in a later section.

of the hospital and consulted with the resident physicians regarding the availability of functional cases. In each instance, the resident physicians were told that the cases desired were those patients who fit the picture of the manifest-anxiety sufferer (on the basis of the definition) and whose complaints were discovered to have no organic basis. Each case submitted by the resident physician was fully discussed and had to meet the stated requirements before being accepted for the study.

2. *The Normal ("N") Group*

Owing to the assorted composition of this group, it is apparent that the group cannot be considered as a control, in the true sense of the term. It is conceivable that among the "normal" population there are some psychotics, prepsychotics, and severe cases of anxiety neuroses. In addition, there is question whether any one is completely free of the trait of anxiety. The control group should therefore be regarded as a group of "normals" which is to be differentiated from the "criterion" group on the grounds that the latter group displayed extreme manifest anxiety.

Particular attention was paid to the matter of representative sampling. As in the case of the criterion sampling, a variety of sources was tapped. For example, the outside or nonhospital "N" group (other groups will be discussed in the following pages) was composed of (a) students from several different college classes, (b) salesmen, and (c) clerical workers.

G. DESIGN OF STUDY

The design of the study embodied the following considerations and procedures:

1. All groups studied for validation of the test were matched, as far as possi-

ble, on the variables of age and education so that these two factors would not materially influence the results of the investigation.

2. The problem of measuring manifest anxiety and not other variables inherent in the situation constituted a major consideration. To compare hospital "MA" cases with the nonhospital (outside) "N" cases alone would not suffice, since the factors of hospitalization could readily affect the measurement. Hence, it became necessary to attempt control of this variable by application of the measuring instrument to a similar group of hospital cases characterized by equal length of stay in the hospital, and by organic illness without "MA" involvement. With such considerations in mind, the following subgroups were composed in order to obtain items which would effect valid differentiation:

- a. Hospital "MA" cases were compared with hospital "N" cases. The first step in the standardization of the test was to develop an adequate number of item-differentiators for these two groups. However, in order to make sure that hospitalization was not exerting undue influence on the results from the "N" cases, the group comparisons described in *b* and *c* below were also made.

- b. Nonhospital "N" cases were compared with the hospital "MA" cases. The purpose here was to eliminate the factor of hospitalization as an influence on the development of anxiety in the "N" group. The factor of item-identity also required consideration. Presuming that the item-differentiators secured in the first investigation were for the most part measuring "MA" and not hospitalization, it would be necessary for the same item-discriminators to effect discrimination of the groups in this second procedure.

- c. Nonhospital "N" cases were com-

pared with hospital "N" cases. The purpose here was to attempt a further control of the factor of hospitalization. If the measuring instrument is a valid one, relatively few items (ideally, none) should discriminate between these two groups.

3. Since certain groups would necessarily contain both sexes, the problem of possibly measuring sex differences rather than "MA" differences also had to be taken into account. A further step in the validation procedures, therefore, was the comparison of a new group of 100 hospital "MA" male patients with a new group of 50 nonhospital female subjects, to determine if sex difference influenced the results of the investigation.

4. A reliability study was performed upon a new group of 100 nonhospital "N" subjects.

5. The reliability study was extended to include those items which were "nearly significant," to determine if the inclusion of these items would increase the reliability of the test.

6. The design further examined the significant items to determine (a) which type of item was most successful in differentiating the "MA" from the non-"MA" cases; and (b) what implications could be derived from the content of the valid items.

7. Finally, the validity of total scores on the test was determined, and norms were established.

H. VALIDATION AND STANDARDIZATION OF THE TEST

The test was validated on a total of 921 subjects.⁵ Of this number, 405 were

used in preliminary investigations (160 criterion cases and 245 control subjects) of both sexes. Since a detailed report of the results of these preliminary studies would heavily overload this paper with statistical data, only occasional reference will be made to these earlier investigations in the course of the present study.

The 616 different subjects comprising the group used in the *final* procedures in the validation of the test are classifiable as follows:

Item-validation studies: Total 366
150 criterion cases
216 control subjects

Reliability studies: Total 100
100 outside (nonhospital) "N" subjects

Sex-difference study: Total 150
(Includes the 100 "N" of the reliability study)

Cross-validation of total test: Total 200
(Includes the 100 "N" of the reliability study)

The procedure in this investigation for the validation of the test differs from the usual method of test validation. In the latter instance, standardization of a test is accomplished by accumulating a number of significant indicators and then applying the completed instrument to a new group of criterion cases for validation of total scores obtained by the instrument. In this investigation, significant items were developed in preliminary item-analysis studies, and *each item* of the completed instrument was validated by comparison of responses in various "MA" vs. "N" groups. This procedure has the advantage of validating every single item of the test, through the establishment of item-identity in all of the validity comparisons made. The fact that a total of 771 subjects was used in the preliminary and final applications of this validation procedure, of which number 310 were "MA" cases, should adequately establish the validity of the instrument *on the basis of constancy of item-differentiation*.

⁵ The hospital "MA" and "N" cases were obtained from the White Memorial Hospital and the Cedars of Lebanon Hospital in Los Angeles; from the Los Angeles County General Hospital; and from the Veterans Administration Hospital, Van Nuys, California.

I. FINAL ITEM-VALIDATION STUDIES

1 AND 2

The purpose in Final Study No. 1 was to sample 100 "N" hospital cases and 150 "MA" hospital cases for further validation of the significant items obtained with similar groups and with similar numbers in the preliminary investigations. This validation study was confined to hospital cases only, so that the effect of hospitalization might "cancel out" in the two groups.

The groups were successfully matched for education and age distributions, there being no significant differences between the standard deviations for age and education in the two groups (t of .03 and t of 1.00).

Of 141 items, 18 showed significant phi values (10 significant at .05 level and 8 significant at .01 level). In all the investigations, both preliminary and final, Guilford's phi-coefficient was employed as a measure of the power of each item to distinguish between the groups compared.

Final Study No. 2 deals with two groups of "N" cases. One of the groups is institutional, consisting of 100 hospital "N" cases, and the other is a group of 116 outside normals.⁶ The purpose here was to determine if the normals in the hospitals responded to the test items similarly to the nonhospital normals.

An attempt was made to match the two groups for age and education. The results indicate that while the groups were matched on the variable of education, a statistically significant difference exists in the variability of age (t 's = .03 and 2.77).

The statistical findings obtained from

the comparisons of the two groups show eleven significant items.

The tentative observation which may be formed on the basis of these statistical comparisons is that the comparison of the normal groups produced fewer discriminating items than the comparisons of the hospital normals with the "MA" hospital cases. This finding is in line with the data of the preliminary studies, indicating that the items tend specially to discriminate the "MA" cases.

J. FINAL ITEM-VALIDATION STUDY 3

In this investigation, the plan was to compare a group of 116 nonhospital "N" subjects with a group of 150 "MA" hospital patients in order to obtain valid item-differentiators. The first step was to attempt to match the groups on the variables of age and education. The two groups were successfully matched in education (no significant difference between means or standard deviations, respectively). Although the groups were not successfully matched in age, the difference does not appear to hold practical significance. Item analysis of the results of the comparisons of the two groups resulted in securing 40 significant phi values.

Table 1 presents comparisons of the results obtained in the final item-analysis studies, as well as related data.

K. EVALUATION: CORRESPONDENCE OF SIGNIFICANT PHI VALUES IN FINAL STUDY 1 AND FINAL STUDY 3

In comparing the significant differentiators of Study 1 with those of Study 3, the following findings were secured:

1. Fourteen of the 18 significant values in Study 1 were also significant in Study 3, indicating that the test was differentiating "MA" cases from cases in the two "N" groups, and also suggesting

⁶ This is a representative group composed of 21 school teachers, 17 students of a class in speech, 48 salesmen, and 30 industrial applicants.

TABLE 1
PHI COEFFICIENTS IN PRESENT STUDIES, FOR ITEMS FOUND SIGNIFICANTLY
DISCRIMINATIVE IN PREVIOUS STUDIES

Final Study 1 (Males)	Final Study 2 (Males)	Final Study 3 (Males)
100 hospital "N" cases	100 hospital "N" cases	116 nonhospital "N" cases
150 "MA" hospital cases	116 nonhospital "N" cases	150 hospital "MA" cases
Significant $\phi = .121$ Very signif. $\phi = .162$.135 .177	.120 .160
Number of Significant Phi's		
18 (10 sig.) (8 very sig.)	11 (3 sig.) (8 very sig.)	40 (12 sig.) (28 very sig.)

that the significant items of the test are relatively immune to hospitalization influences.

2. Since 4 of the 14 items in Study 1 increased their significance from the 5 to the 1 per cent level in Study 3, this may be taken to mean that either the confinement of hospitalization itself or the presence of an organic illness develops some measure of anxiety in the organically ill, but not enough to identify them with the "MA" syndrome.

3. All of the values which were significant at the 1 per cent level in Study 1 were equally significant in Study 3. This evidence of item-identity further establishes the validity of the indicators.

4. Since in Study 3, 28 of the 40 differentiators were at the 1 per cent level of significance as against 8 of 10 in Study 1, the evidence strongly suggests the observation that the test is measuring successive degrees of anxiety, from the nonhospital normals, through the organically ill (hospitalized) normals, to the hospitalized "MA" cases.

L. FINAL STUDY 1 COMPARED WITH FINAL STUDY 2

In comparing Study 1 (hospital "N" and hospital "MA" cases) with Study 2 (hospital "N" and nonhospital "N"),

the following results were secured:

1. Only 4 of the 18 significant items in Study 1 showed correspondence of significance in Study 2, thereby again indicating that only a slight increase in anxiety may be considered to exist in the hospital "N" cases over the nonhospital "N" subjects.

2. Since only 11 indicators were established in Study 2 between the two groups of "N" cases, 18 between the hospital "N" and the hospital "MA," and 40 between the nonhospital "N" and hospital "MA," the results may be regarded as showing the instrument to be a valid measure of "MA," both as to extreme (hospitalized "MA") cases and the middle of the continuum (hospital "N" cases).

M. FINAL STUDIES 2 AND 3

Comparison of results of these two studies (hospital "N" and nonhospital "N" cases, vs. "MA" hospital and nonhospital "N") gave the following results:

1. Nine of the 11 significant indicators in Study 2 are the same items as in Study 3, thereby establishing item-identity.

2. Eight of the 9 items are at the 1 per cent level of significance in both studies.

3. Study 3 gained 31 new significant

indicators when the nonhospital "N" cases were compared with the criterion cases.

N. DISCRIMINATING ITEMS

In all of the item-analysis studies, the type of item in subsection "A" of the test typically maintained its consistency in being the most successful differentiator between the "MA" and the non-"MA" cases. Of the total number of valid items, 32 were of the "A" classification, only one significant indicator was secured from the "B" category, none from sections "C" and "D," while 7 emerged from the "E" classification.

O. RELIABILITY STUDY I

This investigation was based upon a new group of 100 nonhospital "N" subjects that included: 17 salesmen in industry (who were informed that their employment would depend upon their obtaining a satisfactory score on the test), 20 fraternity applicants (who were given the inventory with the statement that their test scores would serve as a basis for evaluating their qualifications for admission into the fraternity), and 63 students in speech, drama, and commerce. The professors of the respective classes gave the tests to their male students with the instructions that the test results would be used by the professor for his own evaluation of the student.

The Kuder-Richardson formula #21 was employed for a determination of the reliability coefficient. The results of this preliminary investigation are as follows: Range 20, Mean 17.04, *SD* 3.94, and Reliability Coefficient, .49.

Since the reliability coefficient was not high,⁷ the K-R formula was applied

⁷ Guilford (8) in discussing the question of reliability has stated that the old standards for reliability and validity need serious re-evaluation. He cites three dramatic instances where three

in another preliminary study with weighted scores.⁸ The results showed that the coefficient based on weights turned out to be essentially the same ($r_{11} = .48$).

P. RELIABILITY STUDY II

Procedure

The design in this reliability study was to determine whether increasing the number of item-differentiators would raise the reliability of the test. Accordingly, a procedure was followed in which all of the phi coefficients were placed in rank order, and the 56 items which had phi values of .09 or over were selected.⁹

The results of this study indicate that including the "nearly significant" phi's raises the reliability coefficient from .49

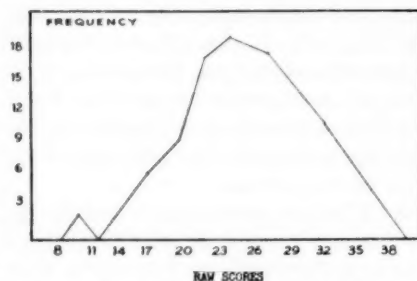


FIG. 1. Distribution of Raw Scores Based on 56 Items: Final Reliability Study, 100 Outside "N" Cases.

in the first reliability study to the relatively high level of .73 in this present

useful tests whose statistics were based on large samples had reliabilities below .40.

⁸ Guilford (6) recommends a simple approximation weight for general use and provides an *abac* for its estimation when the correlation between item and the criterion is the phi coefficient. He derived a formula for the standard error of this weight and presents tables of significant and very significant weights in terms of deviations from the median weight.

⁹ The minimal value of .12 was used in the preceding studies. This "N" group represents the same group of 100 "N" cases used in the preliminary reliability study.

TABLE 2
COMPARISON OF TOTAL SCORES OF 100 HOSPITAL "MA" AND
100 NONHOSPITAL NORMAL CASES
(Total Scores Based on 56 Items)

Group	Range of Scores	Mean	SD
100 "MA" hospital cases	30	29.30	7.29
100 "N" nonhospital cases	20	24.81	6.18
Critical score: 28			
Tests of Significance			
<i>t</i> of Means			3.90
Validity Coefficient ϕ			.32
Equivalent of Pearson r (r_ϕ)			.40*
Chi Square			13.31**

* Since the scores on the tests represent a continuous distribution and the two groups of "N" and "MA" act as a dichotomy and may be considered a point distribution, ϕ becomes equivalent to a Pearson r by the formula $r_\phi = \frac{\phi}{.798}$.

** With one degree of freedom, a chi square of 6.635 indicates significance at the 1 per cent level.

study. Range of scores was 29; the mean score became 24.81; the *SD*, 6.18.

Figure 1 provides a graphic distribution of the scores. It is to be observed that the distribution approximates a normal curve.¹⁰

Q. STUDY OF SEX DIFFERENCE AS A FACTOR IN ANXIETY

The investigation here was concerned with two new groups of nonhospital "N" subjects (100 males and 50 females) selected at random from college students, to determine if the variable of sex difference is a contributing factor to anxiety.

On the basis of the results obtained in this study, it may be stated that no significant difference exists between the scores of the two groups. Sex does not seem to be a determining factor in the formation of anxiety (*t* of means = .89).

¹⁰ The 56 items with significant or "nearly significant" ϕ values are identified in the "MA" test in the appendix.

R. VALIDATION OF TEST: TOTAL SCORES

The purpose of this study was to validate the total scores of the test, using new groups of outside "N" and "MA" cases for the development of a scoring key. The procedures employed were as follows:

1. A new group of 100 "MA" hospital cases was given the test.¹¹
2. The scores were compared with the scores of the 100 nonhospital "N" cases which had been used in the previous reliability study.
3. Fisher's *t* test of significance was applied to the means of the test scores.
4. A cut-off score was established by the graphic method.
5. The ϕ coefficient and chi square were employed for validity studies.

Table 2 presents relevant data.

1. *t* of Means

The data in Table 2 show a very sig-

¹¹ The same procedures were followed by the psychiatrists in the selection of this new group of 100 "MA" cases as characterized the previous psychiatric samplings of the hospital "MA" cases.

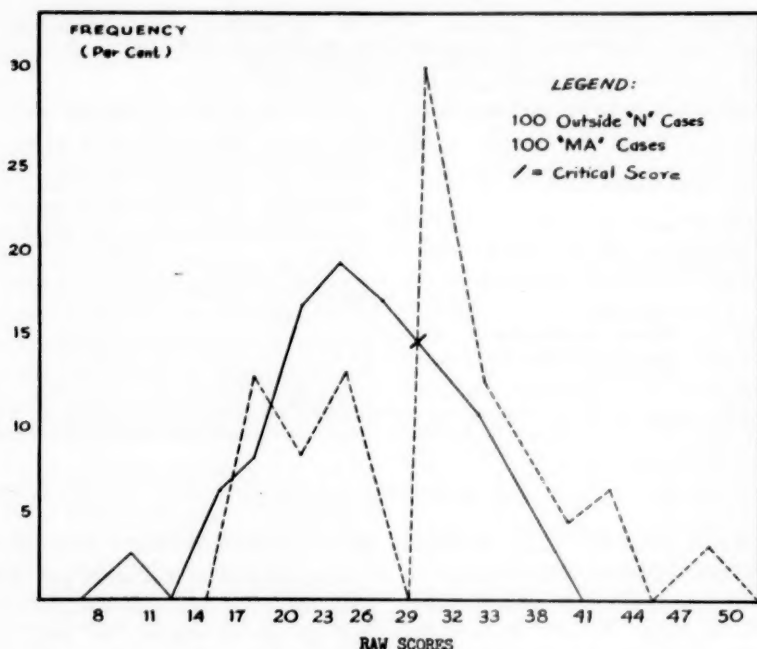


FIG. 2. Distribution of Raw Scores Based on 56 Items: Validation Study, 100 Outside "N" Cases and 100 "MA" Cases.

nificant difference in the mean scores when the "N" group is compared with the "MA" group. This would indicate that there is less than one chance in a hundred that the observed mean difference between "N" and "MA" scores could be produced by chance. The null hypothesis can therefore be rejected with confidence.

Figure 2 shows the distributions of the scores.

2. Critical score

The cut-off score was established by the graphic method of locating the point of intersection of the curves of the distributions of the two groups. The critical score is 28.

3. Validity coefficient "phi" and value of "chi square"

The validity coefficient, Φ , equals .32;

in terms of equivalent Pearson r , this equals .40. For a significance-test, phi was converted into chi square (7, p. 246).¹² As indicated in Table 2, chi square equals 13.31. Since the required figure for significance at the 1 per cent level is 6.635, there is less than one chance in a hundred that the obtained chi square could occur by chance. The result is therefore very significant and the null hypothesis can be rejected with confidence.

S. THE CRITICAL SCORE

The critical score used in separating normals from abnormals cannot be con-

¹² The formula reads $\chi^2 = N\Phi^2$. Guilford states that the standard error for phi is too laborious to be of practical value and he recommends that, for tests of significance, chi square, derived from phi, be employed.

sidered as a clear-cut dividing line. The problem of such separation has always been a difficult matter which has elicited much discussion (2). The term "MA" has been defined in this study as amounts of anxiety observed to be present in a very severe degree. Yet "normals" who possess anxiety in less amounts (below the critical score) may also show disturbed behavior.

The clinical setting, however, demands some dividing line in order to assist in the identification of the "MA" case. To this extent, a critical score becomes valuable.

TABLE 3
"MA" CENTILE NORMS BASED ON 100
NONHOSPITAL "N" CASES

Centile	Score Point	Integral Score
99	38.5	39
95	34.5	35
90	33.0	33
85	32.0	32
80	30.0	30
75	29.0	29
70 (c.s.)*	28.2	28
60	26.7	27
50	25.0	25
40	23.3	23
30	21.5	22
20	19.8	20
10	16.6	17

* c.s. = cut-off score.

In order to complete the development of the scoring key, a percentile table has been constructed, as shown in Table 3. The cut-off score falls at the 70th percentile of nonhospital "N" cases, which means that 30 per cent of the nonhospital "N" population have scores higher than the cut-off scores presented by the test.

T. CONCLUSION AND IMPLICATIONS

The preliminary item-analysis studies, the three final item-validation studies, the two reliability studies, the further

study of sex difference, and the final validation study for the establishment of a cutting score and norms seem to justify the following conclusions:

1. A pencil-and-paper personality test can become a reasonably valid measuring instrument if structured on a projective-identification basis such as the "MA" Test developed in this investigation. The "MA" Test maintains the advantage over other objective scoring instruments through (a) its feature of disguise, the test being ostensibly designed to evaluate the subject's ability to judge (socially) the behavior of other people; (b) elimination of conscious insight into the character and intent of the items, thereby reducing the factor of "faking" and manipulation of the test items; and (c) causing the respondent to project his unconscious identifications for revelation of his personal anxiety characteristics. Inspection of responses to the individual items of the "MA" Test also allows the examiner to locate the significant indicators of manifest anxiety and to estimate the nature and extent of the subject's manifestations of anxiety.

2. The implications resulting from the above findings should prove to be highly significant for clinical psychology. If, as seems likely, anxiety reflects a fundamental, nuclear manifestation in personality disorders, it would seem that future clinical research should be directed towards a more adequate understanding of the effects of fear upon the development of emotional pathology and biofunctional disturbances. Similarly, from the standpoint of psychotherapy, effective reconstruction of the personality would need to take into account adequate modification or eradication of the nuclear pathological elements of fear and anxiety.

3. The securing of a low score on the

"MA" Test should not be regarded as an indicator of absence of anxiety or presence of good emotional adjustment. Empirically, the results suggest the conclusion that scores below 1 SD from the mean may be found to characterize either the psychopaths or the anxiety cases who have developed a "comfort trend" in their behavior. In both these instances, the neurosis manifestations would involve a pathological deficit of normal apprehension and conflict through "blotting out" reality, or through relatively uninhibited libidinal expression.

4. Of particular significance is the relationship of the "MA" (manifest anxiety) nucleus to the various pathological directions of the neurosis disorder, including the intellectual functioning. It

is suggested that future research could take into consideration possible relationships between high "MA" scores and clinically suggestive signs of response patterning on the Wechsler-Bellevue test.

The results of this investigation point in the direction of anxiety being nuclear to various forms of neurosis. The basis for this statement is the fact that all forms of neurotic disorder are represented in the group of criterion cases, who were selected only for "manifest anxiety" (see pages 3-4).¹³

¹³ A subsequent paper based on further research suggests the hypothesis that manifest anxiety is nuclear to the psychoneuroses in general. If still further research bears this out, the significance of the "MA" test will obviously be greatly enhanced.

APPENDIX

THE FREEMAN MA TEST¹

Name Age Sex
 Address Occupation
 Date of Test Status

INSTRUCTIONS

The purpose of this test is to tell how well you are able to judge the behavior of other people. In some questions you are asked merely to check "yes" or "no." In others you are to choose the right answer. Read each question carefully, but do not spend too much time on any one question.

EXAMPLE

Questions

1. A boastful person has a weak character
2. A very fussy person is
 - (a) very hard to get along with
 - (b) no trouble at all

Answers

Yes () No (X)
 a (X) b ()

"A"

QUESTIONS

1. When someone says the wrong thing in conversation, he is bothered by it afterwards.
2. One who often loses his temper feels sorry afterwards.
3. If someone is easily irritated, the reason is he cannot help himself.

ANSWERS²

Phi³

Yes (X) No () .17
 Yes (X) No () .35
 Yes (X) No () .18

¹ Copyright, 1952.

² "X" = response scored positively as an indicator of manifest anxiety ("MA").

³ Only values of phi equal to .09 or greater are listed. Most of these items are statistically significant at the .05 level or better. (Cf. Table 1.)

QUESTIONS

ANSWERS

Phi

4. If a person worries a great deal about his job, the reason is that he is afraid of losing his job.	Yes (X) No ()	.13
5. When a person realizes that he has made a bad mistake, he usually condemns himself for having made that mistake.	Yes () No ()	...
6. Most people believe that a murderer should get the electric chair.	Yes () No ()	...
7. The average person becomes "weak" at the sight of blood.	Yes (X) No ()	.16
8. Looking down from a high building is usually frightening.	Yes (X) No ()	.26
9. It is a very good thing that most people go to the doctor at the first sign of illness or pain.	Yes (X) No ()	.14
10. Most people are unable to appreciate how miserable it is to have trouble sleeping at night.	Yes (X) No ()	.19
11. Going to the blood bank for the first time usually causes a person to be fearful.	Yes (X) No ()	.19
12. A good driver is one who always has on his mind the many possible accidents that can happen to him.	Yes (X) No ()	.17
13. Usually when there is pain in one part of the body, other parts of the body also become affected with pain.	Yes (X) No ()	.10
14. The conscientious person continually keeps thinking of all the other things that have to be done before the day's work is finished.	Yes (X) No ()	.25
15. In order to be successful, a person should always be worried about doing the wrong thing.	Yes (X) No ()	.21
16. If a person is very sensitive, the reason is that he has a fine character.	Yes (X) No ()	.20
17. From the standpoint of desirable health standards, most people are lacking in sufficient cleanliness.	Yes (X) No ()	.34
18. The person who is exceptionally concerned about his future security is a practical person.	Yes (X) No ()	.18
19. The person who always has trouble making a decision wants to be sure he is not going to make a mistake.	Yes (X) No ()	.28
20. Being in big crowds of people will frequently cause a person to experience a feeling of suffocation.	Yes (X) No ()	.47
21. A sensible person always keeps the window blinds down at night to avoid being seen.	Yes (X) No ()	.33
22. People who have very few friends are likely to be very selfish people.	Yes (X) No ()	.21
23. A person who has a great deal of energy is usually a very nervous person.	Yes (X) No ()	.15
24. An honest person will very seldom become a politician.	Yes () No ()	...
25. The person who has a great many fears is usually a coward.	Yes (X) No ()	.14
26. People who reach the top of the ladder in industry usually get there because of some pull.	Yes (X) No ()	.26
27. Many people suffer from extreme fatigue.	Yes () No ()	...
28. Going to the movies is one of the best ways to overcome discouragement.	Yes (X) No ()	.19
29. The difficulty with most relatives is that they are usually very bothersome and jealous.	Yes (X) No ()	.27
30. There are very few people who really understand us well.	Yes () No ()	...
31. It is correct to say that the average person wastes much of his time.	Yes () No ()	...
32. Most people find it particularly annoying to have to wait for a friend who is late for his appointment.	Yes (X) No ()	.11
33. Most of us live under a "nervous tension" because of our "city way" of living.	Yes (X) No ()	.26
34. Usually it is best not to trust anything a salesman says of his product.	Yes (X) No ()	.29
35. The average person finds himself putting heart and soul into his work.	Yes () No ()	...
36. The average person very often checks his work to make sure that he will not make a mistake.	Yes () No ()	...
37. In handling daily duties, one cannot be at ease until the very last thing is out of the way.	Yes (X) No ()	.16
38. As a rule, one does not tell a person about his faults because one does not wish to hurt his feelings.	Yes () No ()	...
39. When something of importance is about to take place, a person usually thinks of the many things that may go wrong.	Yes () No ()	...

QUESTIONS

40. The trouble with most people is that they do not pay enough attention to turning out perfect work.
41. Very few people are sufficiently tidy about their work.
42. Only some people have the exceptional ability to judge whether a person is telling the truth.
43. It is usually very upsetting to have to read in the newspapers about a terrible tragedy that has taken place.
44. It is best to have only a few friends, but good ones.
45. One typically finds that the experiences which were at first feared turned out to have no danger to them at all.
46. Most people find it disturbing to read of the many murders and accidents that are reported in the newspapers.
47. As a rule people's fears concern small silly things rather than large events.
48. In this day and age it is very difficult to find enough time in which to relax.
49. One has always so many things to do that it becomes difficult to avoid being rushed.
50. There are more dishonest than honest people in the world.
51. A conscientious person is one who spends a lot of time thinking about his problems.

ANSWERS

Phi

- Yes () No () ...
- Yes (X) No () .11
- Yes (X) No () .14
- Yes () No () ...
- Yes (X) No () .22
- Yes () No () ...
- Yes (X) No () .23
- Yes () No () ...
- Yes (X) No () .20
- Yes (X) No () .19
- Yes (X) No () .30
- Yes (X) No () .23

"B" (CHOOSE 1 or 2)

52. The person who presents the more undesirable behavior is
(1) the one who criticizes people
(2) the person who lies occasionally
53. The less bothersome person would be
(1) the one who argues a lot
(2) the one who has a weak character
54. Someone has trouble making a decision. The reason is that
(1) he is giving careful thought to both sides of the question
(2) he doesn't trust his judgment
55. The one who would prove to be most undesirable as a friend would be
(1) the person who is very stubborn
(2) the person who lacks education
56. A person would suffer more from
(1) nervousness
(2) rheumatism
57. The person who works rapidly at his job is
(1) a nervous worker
(2) an efficient person
58. A person is the life of the party because
(1) he is always happy
(2) he is usually unhappy
59. The successful person
(1) takes things as they come
(2) worries before each task
60. A person with good character feels sorry for people in trouble
(1) once in a while
(2) all the time
61. People are usually more bothered by
(1) loud noises
(2) headaches
62. People are usually more aggravated by
(1) being disappointed about something
(2) not being able to solve problems
63. The average person attempts to enjoy himself by
(1) usually doing something
(2) usually relaxing

- 1 (X) 2 () .09
- 1 () 2 () ...
- 1 (X) 2 () .13
- 1 (X) 2 () .10
- 1 () 2 () ...
- 1 () 2 () ...
- 1 (X) 2 () .12
- 1 (X) 2 () .11
- 1 () 2 () ...
- 1 () 2 () ...
- 1 () 2 () ...
- 1 () 2 () ...

QUESTIONS

ANSWERS Phi

64. As a rule, when something disturbing takes place
 (1) it does not last long
 (2) usually causes a person to take it to heart
 1 () 2 () ...
65. Of the two, the more objectionable person is the one
 (1) who is dishonest at times
 (2) who talks too much
 1 () 2 () ...
66. A person is constantly active because
 (1) he does not like to waste his time
 (2) he finds it hard to do nothing
 1 (X) 2 () .27
67. A person who wants to satisfy his wishes without much delay is
 (1) an impulsive person
 (2) a very successful person
 1 () 2 () ...
68. John Doe has these two fears. He is more disturbed by
 (1) fear of death through an automobile accident
 (2) fear of dying during the night
 1 (X) 2 () .09
69. Of the two, a person is more embarrassed when
 (1) realizing he was lost in thought while someone was talking to him
 (2) dropping his fork at the dinner table
 1 () 2 () ...
70. People as a rule suffer mostly from
 (1) general aches and pains
 (2) one specific illness
 1 () 2 () ...
71. Forgetfulness usually takes place
 (1) at work
 (2) at occasions when someone's name has to be recalled
 1 () 2 () ...
72. (1) most people are very conscientious
 (2) few people are very conscientious
 1 () 2 () ...
73. People are more frequently disturbed by
 (1) looking down from a tall building
 (2) not having sufficient air in the room
 1 () 2 () ...
- "C" (CHOOSE 1, 2, or 3)
74. When a supervisor is too "bossy" the employee feels that
 (1) the supervisor should be "fired"
 (2) should be lowered to the employee's rank
 (3) should not pay attention to the supervisor
 1 () 2 () 3 () ...
75. A person who prefers to be alone is usually
 (1) an unfriendly person
 (2) likes to mind his own business
 (3) a home lover
 1 () 2 () 3 () ...
76. As far as imagination is concerned, most people have
 (1) just an ordinary amount
 (2) a small amount
 (3) a great amount
 1 () 2 () 3 () ...
77. One usually finds that the average person regards his personal belongings with
 (1) ordinary care
 (2) very careful attention
 (3) not enough care
 1 () 2 () 3 () ...
78. A person usually worries about things because
 (1) he can't help himself
 (2) he gives serious thought to his problems
 (3) tries to avoid doing the wrong thing
 1 () 2 () 3 () ...
79. One of your friends is bothered by these three things. The one that aggravates him the most is
 (1) shyness
 (2) poverty
 (3) irritability
 1 () 2 () 3 () ...
80. Another one of your friends suffers from the following. The most bothersome one is

QUESTIONS

ANSWERS Phi

- (1) sensitiveness
(2) rheumatism
(3) nervousness
1 () 2 () 3 () ...
81. The ideal employee would be one who is
(1) exceptionally careful about details
(2) ordinarily careful
(3) doesn't waste too much time about details
1 () 2 () 3 () ...
82. Judging our friends, we should say that most of them are
(1) exceptionally happy
(2) occasionally happy
(3) never really happy
1 () 2 () 3 () ...
83. In trying to be successful at work, one should
(1) stay overtime and try to get everything done
(2) try not to become discouraged
(3) work conscientiously
1 () 2 () 3 () ...
84. John Doe has three problems. He will first try to overcome
(1) being fearful
(2) being in debt
(3) lack of education
1 () 2 () 3 () ...

"D" (CHOOSE ONE)

85. Someone you know shows these undesirable traits. The most aggravating trait that he should first start to overcome is
(1) telling lies
(2) not being on time
(3) laziness
(4) nervousness
1 () 2 () ...
3 () 4 () ...
86. A person who has all of the following ailments is probably mostly upset by
(1) inability to sleep well
(2) continual head colds
(3) upset stomach
(4) earache
1 () 2 () ...
3 () 4 () ...
87. Of the following four traits, the most undesirable one is
(1) a person who continually criticizes
(2) who is domineering
(3) who shows dishonesty
(4) who does not save any money
1 () 2 () ...
3 () 4 () ...
88. If one of your friends showed the following traits, the most disturbing one would be
(1) his being superstitious
(2) his becoming easily irritable
(3) having feelings of insecurity
(4) having many prejudices
1 () 2 () ...
3 () 4 () ...
89. A person hurts himself mostly from
(1) working too much
(2) frequently losing his temper
(3) too much anxiety
(4) not saving for a "rainy day"
1 () 2 () ...
3 () 4 () ...

"E"

Below is a list of paired characteristics. You are to place an "X" before each one of the pairs which comes the closest to fitting the behavior of the average person.

Here is an example

- () Forgetfulness
(X) Does not read enough

Be sure that you mark one of each set of statements.

Phi

- ... 90. () Tendency to be lost in thought
() Does not read enough

Phi

- ... 91. () Becomes discouraged at times
() Irritable at times
- ... 92. () Inclined towards worry
() Inclined towards fast driving
- ... 93. () Worries about work
() Does not belong to enough organizations
- ... 94. () Has few friends
() Makes mistakes occasionally
- .15 95. (X) Fears looking down from a tall building
() Fears riding in an elevator
- .11 96. (X) Fears losing his job
() Fears making a mistake
- ... 97. () Feels miserable at times
() Feels lazy at times
- ... 98. () Often severely criticizes himself for doing the wrong thing
() Often forgets to remember peoples' names after first meeting them
- .09 99. (X) Not satisfied until every detail is properly handled
() Not satisfied until objectives are reached
- .21 100. (X) Fears another war taking place
() Fears doing the wrong thing
- ... 101. () Tendency to postpone doing things
() Does not take enough time for lunch
- ... 102. () Tendency to overlook details
() Dislikes riding street cars
- ... 103. () Too exacting in work
() Occasional carelessness
- ... 104. () Dislikes being bossed
() Dislikes heavy traffic
- ... 105. () Keeps hurt feelings to himself
() Occasional loss of temper
- ... 106. () Tries to keep from being aggravated
() Tries not to be late
- ... 107. () Finds it difficult to make new friends
() Finds it difficult to save enough money
- ... 108. () Tries hard to be nice to people
() Tries to be well informed on all subjects
- ... 109. () Tendency to look at bad side of future happenings
() A dislike for getting into an argument
- .09 110. (X) Has been called stubborn by others
() Has been criticized for being late for an appointment
- ... 111. () Has been told that he is forgetful
() Has been criticized for being late for an appointment
- .10 112. (X) Often cannot overcome certain problems
() Often changes ideas about things
- .13 113. (X) Wishes people were more tactful
() Wishes people were more sensible
- ... 114. () Wishes he had more time in which to do things
() Wishes people would be more cooperative
- ... 115. () Often regrets saying things
() Often likes to sleep later in the morning
- .13 116. (X) Has habit of being too careful in making a decision
() Has habit of not giving enough thought to old age security.
- .13 117. (X) Has tendency to feel bad when things do not work out as they should
() Has tendency to forget to remember a relative's birthday.
- ... 118. () Likes to attend movies frequently
() Likes to play cards with others
- ... 119. () Believes in leading a practical life
() Believes in good neighbor policy
- .18 120. (X) Always active and on the go

Phi

- () Always interested in enjoying life.
- ... 121. () Dislikes being idle
- () Dislikes Communism
- ... 122. () Eager to know what the next day will bring
- () Eager to know the news of the world
- ... 123. () Wishes there were less unhappiness in this world
- () Wishes parking facilities were better
- ... 124. () Believes nervousness can be overcome
- () Believes nervousness is very difficult to overcome
- ... 125. () Wants to get rid of all personal fears
- () Wants to get rid of all unfinished tasks
- ... 126. () Would like to see Capital and Labor get along
- () Would like to have people be more friendly
- ... 127. () Would like to see more education for everyone
- () Would like less politics in government
- ... 128. () At times has trouble making a decision
- () At times feels a need for rest
- ... 129. () Would wish people would drive more carefully
- () Would wish a work week would consist of three days
- ... 130. () Would regard lack of patience as being a form of nervousness
- () Would say lack of patience is due to over-anxiousness in wanting to get a thing done
- ... 131. () Would like to enjoy a good night's sleep
- () Would like to take a vacation
- ... 132. () Would always like to have a healthy heart
- () Would always like to be happy
- 15 133. (X) Would like to see a better world in which to live
- () Would like to see a reduction of taxes
- ... 134. () Believes people need to understand him better
- () Believes people need to develop more education
- ... 135. () Most people are careless in doing their work
- () Most people need more hobbies
- ... 136. () The average person works leisurely
- () The average person finds a day's work is usually tiring
- ... 137. () The average employer fully fails to appreciate his employee
- () The average employer doesn't take time to learn what his employee accomplishes
- 10 138. (X) Wishes people's criticisms wouldn't bother him so much
- () Wishes industry were capable of more production
- ... 139. () Finds that when he hates someone it doesn't last long
- () Many people could develop better judgment
- ... 140. () Whenever a person is restless he seeks to overcome it
- () Whenever someone has cheated, he is distrusted thereafter
- ... 141. () Finds that usually when he fears something is going to happen, it does not take place
- () Finds that he dislikes careless drivers.

REFERENCES

1. BABCOCK, H. A. A neurosis? Or neurotic behavior? *J. Psychol.*, 1944, 17, 61-74.
2. ELLIS, A. The validity of personality questionnaires. *Psychol. Bull.*, 1946, 43, 385-441.
3. FREEMAN, M. J. The standardization of a psychosomatic test: validation of a psychosomatic inventory. *J. Pers.*, 1950, 19, 229-244.
4. GUILFORD, J. P. *Psychometric methods*. New York: McGraw-Hill, 1936.
5. GUILFORD, J. P. The phi co-efficient and chi square as indices of item validity. *Psychometrika*, 1941, 6, 11-19.
6. GUILFORD, J. P. A simple scoring weight for test items and its reliability. *Psychometrika*, 1941, 6, 367-374.
7. GUILFORD, J. P. *Fundamental statistics in psychology and education*. New York: McGraw-Hill, 1942.
8. GUILFORD, J. P. New standards for test evaluation. *Educ. psychol. Measmt.*, 1946, 4, 427-438.
9. GUILFORD, J. P., & GUILFORD, RUTH B. An analysis of the factors in a typical test of introversion-extroversion. *J. abnorm. soc. Psychol.*, 1934, 28, 377-399.
10. KUDER, G. F., & RICHARDSON, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, 2, 151-160.
11. KUZNETS, G. An analysis of Bernreuter's personality inventory. *Psychol. Bull.*, 1934, 31, 585.
12. LANDIS, C. Questionnaires and the study of personality. *J. nerv. ment. Dis.*, 1936, 83, 125-134.
13. LANDIS, C., & KATZ, S. E. The validity of certain questions which purport to measure neurotic tendencies. *J. appl. Psychol.*, 1934, 18, 343-356.
14. LANDIS, C., ZUBIN, J., & KATZ, S. E. Empirical evaluation of three personality adjustment inventories. *J. educ. Psychol.*, 1935, 26, 321-330.
15. LORGE, I. Personality traits by fiat, I and II. *J. educ. Psychol.*, 1935, 26, 273-278; 427-434; 652-654.
16. MALAMUD, D. C. Objective measurement of clinical status in psychopathological research. *Psychol. Bull.*, 1943, 3, 240-259.
17. MAY, R. *The meaning of anxiety*. New York: Ronald, 1950.
18. MEEHL, P. An investigation of a general normality or control factor in personality testing. *Psychol. Monogr.*, 1945, 59, No. 4 (Whole No. 274).
19. RULON, P. J. A simplified procedure for determining the reliability of a test by split-half. *Harvard educ. Rev.*, 1939, 9, 99-103.
20. SULLIVAN, H. S. *Conceptions of modern psychiatry*. Washington, D.C.: William Alan-son White Psychiatric Foundation, 1947.
21. WHITE, R. W. *The abnormal personality*. New York: Ronald, 1948.

(Accepted for publication September 25, 1952.)